

## WHAT IS CLAIMED IS:

1. A method for utterance verification comprising the steps of:

(A) extracting a sequence of feature vectors from an input speech;

(B) inputting the sequence of feature vectors to a speech recognizer for  
5 obtaining at least one candidate string;

(C) segmenting the input speech into at least one speech segment according to the content of candidate string, which comprises individual recognition units, wherein each speech segment corresponds to a recognition unit and each recognition unit corresponds to a verification unit;

10 (D) generating a sequence of verification feature vectors for each speech segment according to the sequence of feature vectors of the speech segment, wherein the verification feature vectors are generated by normalizing the feature vectors using the normalization parameters of the verification unit corresponding to the speech segment;

15 (E) utilizing a verification-unit corresponded classifier for each speech segment to calculate the verification score, where the sequence of verification feature vectors of the speech segment is used as the input of the classifier;

(F) combining the verification scores of all speech segments for  
20 obtaining an utterance verification score of the candidate string; and

(G) comparing the utterance verification score of the candidate string with a predetermined threshold so as to accept the candidate string if the utterance verification score is larger than the predetermined threshold.

2. The method as claimed in claim 1, wherein in step (D), the

normalization parameters of the verification unit are the means and the standard deviations of the feature vectors corresponding to the verification unit in training data, and these parameters are calculated in advance.

3. The method as claimed in claim 1, wherein in step (E), the classifier  
5 is a neural network, and the neural network is an MLP (multi-layer perceptron).

4. The method as claimed in claim 3, wherein the MLP is used to calculate the verification score by inputting the verification feature vector and performing the feed-forward processing, the verification score of a  
10 speech segment is the mean of the verification scores of the sequence of verification feature vectors corresponding to the speech segment.

5. The method as claimed in claim 3, wherein the MLP is trained by using an error back-propagation algorithm to reduce the mean square error between the verification score output of the MLP and the target value.

15 6. The method as claimed in claim 5, wherein the MLP corresponding to the verification unit is trained by inputting the sequences of verification feature vectors of the speech segments corresponding to the verification unit and the sequences of verification feature vectors of the speech segments not corresponding to the verification unit.

20 7. The method as claimed in claim 6, wherein the target value is 1 if the speech segment corresponds to the verification unit and which is 0 if the speech segment does not correspond to the verification unit.

8. The method as claimed in claim 1, wherein in step (F), the utterance verification score of the candidate string is the mean of the verification

scores of the speech segments in the input speech.

9. The method as claimed in claim 1, wherein the input speech is corrupted by noise.

10. The method as claimed in claim 6, wherein the speech segments  
5 used for training are corrupted by noise.

11. A system for utterance verification comprising:

a feature vector extraction module for extracting a sequence of feature  
vectors from an input speech;

a speech recognition module for obtaining at least one candidate string  
10 by inputting the sequence of feature vectors;

a speech segmentation module for segmenting the input speech into at  
least one speech segment according to the content of candidate string,  
which comprises individual recognition units, wherein each speech  
segment corresponds to a recognition unit and each recognition unit  
15 corresponds to a verification unit;

a verification feature vector generation module for generating a  
sequence of verification feature vectors for each speech segment according  
to the sequence of feature vectors of the speech segment, wherein the  
verification feature vectors are generated by normalizing the feature vectors  
20 using the normalization parameters of the verification unit corresponding to  
the speech segment;

a verification score calculation module for utilizing a verification-unit  
corresponded classifier for each speech segment to calculate the  
verification score, where the sequence of verification feature vectors of the

speech segment is used as the input of the classifier;

a verification score combination module for combining the verification scores of all speech segments for obtaining an utterance verification score of the candidate string; and

- 5 a decision module for comparing the utterance verification score of the candidate string with a predetermined threshold so as to accept the candidate string if the utterance verification score is larger than the predetermined threshold.

12. The system as claimed in claim 11, wherein in the verification  
10 feature vector generation module, the normalization parameters of the verification unit are the means and the standard deviations of the feature vectors corresponding to the verification unit in training data, and these parameters are calculated in advance.

13. The system as claimed in claim 11, wherein the classifier is a neural  
15 network, and the neural network is an MLP (multi-layer perceptron).

14. The system as claimed in claim 13, wherein the MLP is used to calculate the verification score by inputting the verification feature vector and performing the feed-forward processing, the verification score of a speech segment is the mean of the verification scores of the sequence of  
20 verification feature vectors corresponding to the speech segment.

15. The system as claimed in claim 13, wherein the MLP is trained by using an error back-propagation algorithm to reduce the mean square error between the verification score output of the MLP and the target value.

16. The system as claimed in claim 15, wherein the MLP corresponding

to the verification unit is trained by inputting the sequences of verification feature vectors of the speech segments corresponding to the verification unit and the sequences of verification feature vectors of the speech segments not corresponding to the verification unit.

- 5        17. The system as claimed in claim 16, wherein the target value is 1 if the speech segment corresponds to the verification unit and which is 0 if the speech segment does not correspond to the verification unit.

18. The system as claimed in claim 11, wherein in the verification score combination module, the utterance verification score of the candidate string  
10 is the mean of the verification scores of the speech segments in the input speech.

19. The system as claimed in claim 11, wherein the input speech is corrupted by noise.

20. The system as claimed in claim 16, wherein the speech segments  
15 used for training are corrupted by noise.